



Mikroekonomický scoringový model úpadku českých podniků

Jiří VALECKÝ, Eva SLIVKOVÁ, VŠB-TU Ostravaⁱ

Abstract

The paper is devoted to the proposing a scoring model of firm's bankruptcy on the basis of logistic regression which could be used in the purpose of classifying the good and bad firms in the Czech Republic. Firstly, the general foundation of logistic regression is described and then the general procedure of model-building strategy including the statistical verification and in-sample validation are explained. Then, we have built the empirical scoring model using results of financial analysis carried out over 400 Czech firms in 2008 and we select such financial ratios with the most substantial prediction capability, i.e. return on assets, current and long-term indebtedness and cash ratio. Our proposed model is fully statistical verified and we also perform the in-sample validation via classification table and ROC curve. Finally, we compare our model with others studies made in this field.

Keywords

Classification table, default probability, fractional polynomial, logistic regression, logit, ROC curve, scoring.

JEL Classification: C25 G33

ⁱ Department of Finance, Faculty of Economics, VŠB-Technical University of Ostrava, Sokolská 33, 701 21 Ostrava, Czech Republic.

jiri.valecky@vsb.cz (corresponding author)

This paper was solved within the project SP2011/38 Estimating and forecasting individual risk in financial institutions. The authors also wish to express his gratitude to three anonymous referees, whose valuable remarks helped to improve the quality of the paper.

1. Úvod

Predikční modely jsou dnes velmi oblíbeným nástrojem pro predikci téměř čehokoli. Jejich vlastnosti se využívá v různých oblastech společenského života, zejména v lékařství, sociologii, politice, marketingovém výzkumu, ekonomice apod. Ve finanční oblasti poté zastávají důležitou a nezastupitelnou roli při hodnocení bonity a finančního zdraví podniků, a to zejména žádá-li tento subjekt o poskytnutí úvěru v některé z finančních institucí. Právě na potenciálně špatné klienty by měly správně fungující predikční modely včas upozornit, aby nedošlo u investorů a věřitelů k přílišným finančním ztrátám.

V hodnocení určitého subjektu lze rozlišovat rating a scoring. Oba přístupy jsou si podobné a bývají často mylně zaměňovány, avšak rating je mnohem složitější

na zpracování, neboť je zapotřebí větší množství kvalitních vstupních dat.

Rating představuje komplexní hodnocení určitého subjektu s cílem odhadnout jeho současnou a budoucí schopnost splácet včas a řádně své závazky vyplývající z emise cenných papírů nebo i jiné závazky, ke kterým se subjekt zavázal. Náročnost provedení spočívá především v hodnocení velkého množství informací, a to jak interních, tak externích. Na základě provedené fundamentální analýzy a vyhodnocení veškerých známých rizik je pak agenturou udělena ratingová známka, viz Sedláček (2007).

Naproti tomu jsou scoringové modely, v užším slova smyslu, založeny na kvantitativní finanční analýze dat, která hodnocený subjekt předkládá za minulá hospodářská období. Kvalitativní data mohou

být rovněž použita, avšak musí být nejdříve dle určitých obecně platných pravidel transformována na data kvantitativní. Tento předpoklad pak snižuje možnost subjektivního úsudku o hodnoceném podniku. Scoring se používá zejména ve finančním sektoru a pojišťovnictví, kde slouží k detekci pojišťovacích podvodů. V bankovníctví pak scoring slouží k určení klientů, u kterých hrozí riziko, že by přestali splácet úvěr, který jim byl poskytnut. Nejpoužívanějším scoringem je tedy kreditní scoring, který předvídá, zda klient neztratí svou schopnost splácet úvěr.

Za scoringové modely jsou považovány především logitové a probitové modely aplikované na podniky, jež se objevují v prvních pracích, např. Ohlson (1980). Jejimi předchůdci však byly jednoduché klasifikační testy, např. Beaver (1967), nebo modely založené na vícerozměrné diskriminační analýze, např. Altman (1968), nebo u nás známý index IN, Neumaierová a Neumaier (2005). Obecné srovnání scoringových modelů a diskriminační analýzy lze nalézt v pracích McFaddena (1976) nebo Gurného a Gurného (2010b). Souhrnný přehled modelů a přístupů lze nalézt například v práci Altmana a Hotchkisse (2006) či v citačním přehledu Jakubíka a Teplého (2011). V české literatuře jsou mikroekonomické logitové a probitové modely aplikovány zejména na české podniky, Jakubík a Teplý (2008, 2011), nebo na české banky odvozené z amerického bankovního trhu, např. Gurný a Gurný (2010a).

Za obecný problém mikroekonomických scoringových modelů lze považovat sběr dat. Jsme toho názoru, že u mikroekonomických modelů by sběr dat měl být prováděn za kratší časové období v podobě několika let v rámci jednoho stavu ekonomiky. Na jednu stranu lze očekávat nestabilitu odhadnutých koeficientů (v rámci celého ekonomického cyklu), na druhou stranu v delším časovém období nejsou pravděpodobnostmi úpadku od sebe odlišeny dvě stejné firmy v různých fázích ekonomického cyklu, neboť pro všechny fáze ekonomiky jsou stejné koeficienty modelu. Pro zahrnutí celého ekonomického cyklu je vhodné použít vhodnější makroekonomické scoringové modely, jež zahrnují pouze makroekonomické determinanty, např. Jakubík (2007), nebo kombinaci mikro- a makroekonomických determinant, např. Gurný (2011). O makroekonomické determinanty bylo rozšířeno i Altmanovo Z-skóre a byl vytvořen tzv. Z-metrics, viz Altman a kol. (2010). V případě nemožnosti pozorování ekonomického cyklu lze použít i režimové logitové modely na bázi skrytých Markovových řetězců. Za další nedostatek logitových modelů může být chápána jejich linearita, nerespektující měnící se vliv determinantu na výsledné skóre (a následně na pravděpodobnost úpadku).

Cílem tohoto příspěvku je vytvoření mikroekonomického modelu úpadku na bázi logistické regrese pro české firmy. Z důvodu nedostatku dat je tento odhad proveden pouze na malém vzorku z roku 2008. Při sestavení modelu je však ověřena linearita logitu a tato je následně ošetřena kategorizací spojitých mikroekonomických determinantů. Model je v rámci použitého vzorku statisticky verifikován a rovněž je vyhodnocena jeho klasifikační (diskriminační) schopnost v roce 2009. V druhé části jsou nejprve položena nezbytná teoreticko-metodologická východiska. Je zde popsána logistická regrese a poté jsou vysvětleny obecné postupy stavby, verifikace a hodnocení sestaveného modelu. Ve třetí části je poté sestaven a odhadnut empirický model na reálných finančních datech 400 českých podniků. Tento model je zde plně verifikován a vyhodnocen pomocí ROC křivky. V poslední, čtvrté části jsou shrnuty a vyhodnoceny získané výsledky.

2. Logistická regrese

Podstatou scoringového modelu je kvantifikace vztahu mezi závislou vysvětlovanou proměnnou (dichotomickou či multinomickou) a nezávislými vysvětlujícími proměnnými, přičemž cílem je analyzovat vliv vysvětlujících proměnných na proměnnou vysvětlovanou tak, abychom pro objekt s neznámou hodnotou vysvětlované proměnné mohli tuto hodnotu odhadnout. Ke klasifikaci vztahu mezi vysvětlovanou a vysvětlujícími veličinami a pro odhad budoucích hodnot závislé proměnné se používají různé metody, např. diskriminační analýza, logistická regrese, klasifikační stromy, neuronové sítě a další.

Vzhledem k dichotomické vysvětlované proměnné nelze tedy při konstrukci logistického regresního modelu vycházet z klasického lineárního modelu, ale je nutné vyjít ze zobecněného lineárního modelu (*Generalized Linear Model*), neboť zde nejsou splněny některé z podmínek pro lineární regresi (např. předpoklad homoskedasticity nebo normálního rozdělení náhodné složky). Předpokládáme, že pravděpodobnost úpadku i -tého podniku $P_i = P(Y_i = 1)$ na základě jeho charakteristického vektoru \mathbf{x}_i lze vyjádřit funkcí $F(\boldsymbol{\beta}; \mathbf{x}_i)$, jež je monotónně rostoucí $F'(\boldsymbol{\beta}; \mathbf{x}_i) \geq 0$ a má definiční obor $(-\infty; +\infty)$ a obor hodnot $(0, 1)$. Tedy platí, že $F(-\infty) = 0$ a $F(+\infty) = 1$. Funkci pravděpodobnosti úpadku lze poté psát v obecném tvaru

$$P_i = F(\boldsymbol{\beta}; \mathbf{x}_i), \quad (1)$$

kde $\boldsymbol{\beta}$ je vektor parametrů $(\beta_0, \beta_1, \dots, \beta_K)$ a \mathbf{x}_i je charakteristický vektor i -tého subjektu obsahující K prvků (determinantů), tedy $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{iK})$.

Tyto zmíněné vlastnosti jsou splněny mj. kumulativní distribuční funkcí logistického rozdělení ve tvaru

$$P_i = P(Y_i = 1) = F(\beta; \mathbf{x}_i) = \frac{e^{\beta' \mathbf{x}_i}}{1 + e^{\beta' \mathbf{x}_i}}, \quad (2)$$

ježž lze zároveň označit jako funkci pravděpodobnosti úpadku. Naproti tomu lze pravděpodobnost, že subjekt nebude mít finanční problémy, vyjádřit ve tvaru

$$1 - P_i = P(Y_i = 0) = 1 - F(\beta; \mathbf{x}_i) = \frac{1}{1 + e^{\beta' \mathbf{x}_i}}. \quad (3)$$

Dále definujeme podíl pravděpodobností (2) a (3), jenž je znám také jako šance (*odds*), ve tvaru

$$\frac{\pi}{1 - \pi} = \frac{P(Y_i = 1)}{P(Y_i = 0)} = e^{\beta' \mathbf{x}_i}, \quad (4)$$

a dále logaritmicou transformací (log-odds neboli logit) vztahu (4)

$$\ln \left[\frac{\pi}{1 - \pi} \right] = \beta' \mathbf{x}_i = g(\mathbf{x}_i), \quad (5)$$

kde $\ln \frac{\pi}{\pi - 1}$ vyjadřuje logit (logaritmus šance) neboli onu transformující funkci. Protože $\pi \in (0, 1)$, podíl $\frac{\pi}{\pi - 1}$ může nabýt jakékoli nezáporné hodnoty, logit (5) pak může nabýt jakékoli reálné hodnoty.

Vztah uvedený v rovnici (2) však nemusí být vždy lineární, a to zvláště tehdy, jsou-li uvažovány spojitě determinanty. Vztah mezi tímto determinantem a logitem může být vyjádřen skokovou či nelineární funkcí. V takovém případě vede zanedbání této skutečnosti k nepřesnému odhadu logitu a posléze pravděpodobnosti úpadku (či dosaženého skóre). Tento problém lze řešit pomocí částečných polynomů (*fractional polynomial*), dle kterého je vztah logitu a determinantu aproximován pomocí polynomu řádu J s mocninami řádu p_j pro $j = 1, \dots, J$, tedy

$$g(\mathbf{x}) = \beta_0 + \sum_{j=1}^J \beta_j F_j(\mathbf{x}_1) + \beta_2 x_2 + \dots + \beta_K x_K, \quad (6)$$

kde $F_j(\mathbf{x}_1)$ představuje funkci částečného polynomu řádu j , přičemž pro $j = 1$ platí $F_1(\mathbf{x}_1) = x_1^{p_1}$ a pro $j > 1$

$$F_j(\mathbf{x}_1) = \begin{cases} x_1^{p_j}, & p_j \neq p_{j-1}, \\ F_{j-1}(\mathbf{x}_1) \ln(x_1), & p_j = p_{j-1}. \end{cases} \quad (7)$$

Parametry p_j mohou nabývat jakéhokoliv reálného čísla. Pokud však platí, že $p_j = 0$, pak $x_1^0 \Rightarrow \ln(x_1)$. Obecně je volba mocnin, dle kterých se má vztah aproximovat, omezena na předdefinovanou

množinu $S \in \{-2; -1; 0; 5; 0; 0; 5; 1; 2; 3\}$, viz Royston a Altman (1994).

Linearitu logitu (5) lze rovněž zhodnotit graficky, a to pomocí vyhlazeného logitu metodou *lowess* (*Locally Weighted Scatterplot Smoothing*) pro různé úrovně determinantu. Dále je taktéž možné spojitou veličinu kategorizovat pomocí kvantilů a odhadnout koeficienty pro každou úroveň. Tato nelinearita je poté identifikována, pokud se parametry pro každou kategorii výrazně liší. Rovněž lze použít některý ze statistických testů, např. Box-Tidwell nebo Tukey-Pregibon test.

2.1 Odhad parametrů modelu

Parametry logitu jsou odhádovány především metodou maximální věrohodnosti. Jejím základem je maximalizace věrohodnostní funkce, následně je určen logaritmus této funkce, stanoví se parciální derivace logaritmu funkce a anulováním se získá soustava věrohodnostních rovnic, ze kterých se po vyřešení získají maximálně věrohodné odhady neznámých parametrů, více viz Hebák a kol. (2004).

V případě logistické regrese s binomickou vysvětlovanou proměnnou a spojitými vysvětlujícími veličinami jsou jednotlivá pozorování nezávislá a věrohodnostní funkci můžeme zapsat jako součin pravděpodobností jednotlivých pozorování, viz Hosmer a Lemeshow (2000),

$$L(\beta) = \prod_{i=1}^N \pi_i^{y_i} (1 - \pi_i)^{1 - y_i}, \quad (8)$$

kde N je počet sledovaných firem.

Dosadíme-li do rovnice (8) již odvozený vzorec pro logistickou regresi (2) a následně jej upravíme, získáme následující tvar věrohodnostní funkce

$$L(\beta) = \prod_{i=1}^N \frac{(e^{\beta' \mathbf{x}_i})^{y_i}}{1 + e^{\beta' \mathbf{x}_i}}. \quad (9)$$

Odhad parametrů metodou maximální věrohodnosti je získán maximalizací logaritmu rovnice (9) ve tvaru

$$\begin{aligned} L(\beta) &= \ln l(\beta) \\ &= \sum_{i=1}^N Y_i \cdot \ln(\pi(\mathbf{x}_i)) + (1 - Y_i) \cdot \ln(1 - \pi(\mathbf{x}_i)) \end{aligned} \quad (10)$$

za podmínek

$$\frac{\partial L(\beta)}{\partial \beta_j} = 0 \quad \text{pro } j = 1, \dots, K. \quad (11)$$

Soustava věrohodnostních rovnic (11) je však nelineární v parametrech a vyžaduje speciální metody k vyřešení. K odhadu parametrů je potřeba zvolit některý z iteračních postupů, např. Newton-Raphsonovu metodu. Princip metody spočívá v maximalizaci aproximace logaritmu věrohodnostní

funkce v okolí počátečního odhadu pomocí prvních tří členů Taylorova rozvoje, viz Pecáková (2007). V každém iteračním kroku je navíc spočtena kovarianční matice odhadu parametrů, jejíž prvky z hlavní diagonály (rozptyly) lze použít při konstrukci směrodatných chyb bodových odhadů i intervalů spolehlivosti.

2.2 Statistická verifikace modelu

Dalším krokem po odhadu parametrů je statistická verifikace modelu. V prvé řadě je nutné otestovat významnost zvolených determinantů, a tedy odhadnutých koeficientů. Pro tento účel lze použít test založený na Waldově statistice, jenž porovnává odhadnutou hodnotu parametru a jeho směrodatnou odchylku

$$W_j = \frac{\hat{\beta}_j}{\hat{s}(\hat{\beta}_j)}, \quad (12)$$

kde $\hat{\beta}_j$ značí odhadnutou hodnotu parametru a $\hat{s}(\hat{\beta}_j)$ je směrodatná odchylka odhadnutého parametru.

Waldova statistika má asymptoticky normované normální rozdělení a slouží jako testové kritérium pro ověření hypotézy, že parametr $\beta_j = 0$. Tento test je rovněž vhodný pro prvotní nástřel, zda zařadit determinant do modelu.

Jiný způsob testování významnosti koeficientů uvádí Hosmer a Lemeshow (2000). Jejich požadavkem je, aby model obsahující danou proměnnou, poskytoval o vysvětlované veličině více informací než model, který tuto proměnnou neobsahuje. Kvalita modelu je posuzována na základě porovnání napozorovaných hodnot závisle proměnné s jejími teoretickými hodnotami vypočtenými modelem. Test se posuzuje na základě poměru dvou věrohodnostních funkcí (8) pomocí zjištěných a vypočítaných hodnot, tzv. test věrohodnostním poměrem. Vzorec má následující tvar

$$D = -2 \ln \left[\frac{\prod_{i=1}^N \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{1-y_i}}{\prod_{i=1}^N \pi_i^{y_i} (1 - \pi_i)^{1-y_i}} \right] \quad (13)$$

a po úpravě

$$D = -2 \sum_{i=1}^N \left[y_i \ln \left(\frac{\hat{\pi}_i}{\pi_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - \pi_i} \right) \right], \quad (14)$$

kde D je věrohodnostní poměr a v logistické regresi má stejnou vypovídací schopnost jako reziduální součet čtverců v případě lineární regrese. O statistické významnosti testované proměnné je pak rozhodnuto na základě statistické významnosti rozdílu věrohodnostních poměrů dvou modelů, z nichž jeden obsahuje a druhý neobsahuje testovanou veličinu, tedy

$$G = D_0 - D_1, \quad (15)$$

kde D_0 a D_1 jsou věrohodnostní poměry modelu neobsahující a obsahující determinanty x_k . Pro dva různé modely s počtem parametrů p_1 a p_0 , přičemž $p_0 > p_1$ má statistika rozdělení χ^2 s $p_0 - p_1$ stupni volnosti, viz Hosmer a Lemeshow (2000). Porovnáním vypočtené hodnoty se následně rozhodne o statisticky významném vlivu daných veličin na výskyt úpadku společnosti.

2.3 Stavba modelu

Predikční model pomocí logistické regrese vytváříme postupně. Dle Hosmera a Lemeshowa (2000) je vhodné začít jednofaktorovou analýzou, což znamená otestovat každý determinant vůči výskytu úpadku zvlášť. Máme-li jediný spojitý determinant, je k otestování statistické významnosti vhodný již výše uvedený Waldův test. V případě, že vysvětlující determinanty jsou nominální, ordinální či spojitě s malým počtem hodnot, využívá se k tomuto otestování kontingenční tabulka s k úrovní této veličiny. K ověření shody kvality těsnosti daných a nalezených četností se využívá Pearsonův χ^2 test dobré shody s $k-1$ stupni volnosti.

Po prvotní selekci staticky nevýznamných determinant přecházíme k hledání modelu s více veličinami. K tomu lze použít metodu postupné regrese neboli *stepwise regrese*. Princip metody spočívá v testování významnosti nezávislých veličin po přidání další veličiny do modelu. Jednotlivé proměnné se totiž mohou vzájemně ovlivňovat, jsou mezi nimi interakce, což může mít za následek statistickou nevýznamnost některých z předchozích zahrnutých determinantů po přidání nového determinantu do modelu. Hodnocení neprobíhá na tradiční hladině významnosti 5 % (tzv. p-hodnotě), ale na hladině 25 % a to z důvodu, aby nedošlo k opomenutí některé z významnějších proměnných. Konečné rozhodnutí o zařazení některého z parametrů do modelu má však analytik, který model vytváří, neboť z věcného hlediska může mít zahrnutí dané veličiny podstatný význam k popsání zkoumané problematiky.

Rovněž by měla být ověřena linearita logitu (5), a to některým ze způsobů popsaných výše. Dalším krokem je identifikace interakcí mezi jednotlivými determinanty. Přestože jejich identifikace je relativně snadná, modelování a interpretace těchto interakcí je značně problematická, blíže viz např. Ai a Norton (2003), Greene (2010) a další. Nakonec je nutné provést statistickou verifikaci modelu pomocí testů, viz část 2.2. Lze rovněž aplikovat detailní analýzu reziduí.

2.4 Hodnocení výsledného modelu

Pro hodnocení kvality scoringového modelu je vždy rozhodujícím kritériem cíl analýzy, a tím i použitelnost výsledků. Správný model by měl co nejlépe vyhovovat nejen vybraným datům, ale také populaci, ze které výběr pochází. Logistický regresní model můžeme hodnotit podle toho, nakolik je schopen na základě vybraných prvků rozlišovat, u kterých firem úpadek nastane a u kterých nikoli, nebo jej můžeme hodnotit podle toho, nakolik se pro určité kombinace determinant shodují zjištěné a očekávané četnosti úpadku, viz Hebák a kol. (2007).

Nástrojů pro hodnocení kvality výsledného modelu je celá řada, zmíníme dvě základní metody, tedy klasifikační tabulku a křivku ROC.

Klasifikační tabulka se používá pro znázornění schopnosti regresního modelu správně klasifikovat jednotky. Tabulka obsahuje četnosti správně a mylně zařazených firem. Úspěšnost modelu se následně posuzuje podle podílu správně klasifikovaných (nalezneme je na hlavní diagonále tabulky) k celkovému počtu, neboli $(n_{11} + n_{22})/N$, kde n_{11} , n_{22} je počet správně predikovaných firem v úpadku a bez finančních potíží a N je celkový počet firem.

Druhou zmiňovanou metodou k hodnocení diskriminační schopnosti modelu je křivka ROC (*Receiver Operating Characteristic Curve*). Základem této grafické metody je jednotkový čtverec, jehož svislá osa představuje tzv. senzitivitu a vodorovná osa značí hodnotu 1-specificita. Senzitivita vyjadřuje procento firem, které byly správně označeny vzhledem k úpadku společnosti. Vodorovná osa zaznamenává procento firem, u nichž došlo ke špatné klasifikaci. Křivka ROC pak představuje kombinaci hodnot senzitivity a specifity. Nastat mohou dvě extrémní situace. V případě, že křivka ROC má tvar diagonály, model postrádá jakoukoli diskriminační schopnost a není vhodný pro predikci. V druhém extrémním případě splývá křivka s levým horním rohem, což znamená absolutní shodu skutečných a predikovaných hodnot. Model se tak stává stoprocentně úspěšným a vhodným k predikci.

S křivkou ROC přímo souvisí i ukazatel AUC (*Area Under ROC Curve*) neboli plocha pod křivkou ROC. Ukazatel je vhodný k porovnání dvou či více ROC křivek, neboť je převádí na jednu veličinu umožňující jejich komparaci. Ukazatel AUC může nabýt hodnot od 0,5 do 1, přičemž vyšší hodnota znamená přesnější predikční model.

K hodnocení logistického regresního modelu se dále užívají metody založené na modifikaci testu věrohodnostním poměrem, které již zohledňují počet vysvětlujících proměnných. Jde například o Akaikeho

informační kritérium, Goodmanův index nebo Bayesovské kritérium.

3. Mikroekonomický scoringový model úpadku českých firem

Pro účely sestavení empirického modelu na bázi logistické regrese byly z databáze MagnusWeb získány finanční výkazy 400 firem (pouze společnosti s ručením omezeným a akciové společnosti, případně bytová družstva) za hospodářský rok 2008. Z tohoto počtu bylo 200 prosperujících firem, které ke konci roku 2009 neměly finanční problémy. Druhou polovinu tvořily firmy, které se během roku 2009 ocitly ve finanční tísní (prohlášení konkurzu, zahájení insolventního řízení, prohlášení úpadku podniku apod.).

Poté byla provedena finanční analýza firem. Pro tento účel bylo vybráno 19 poměrových ukazatelů finanční analýzy, viz tabulka 1. Některé ukazatele záměrně vybrány nebyly. Například byla vynechána rentabilita vlastního kapitálu, neboť tento ukazatel nabýval hodnot, jež nereflektovaly skutečný stav.¹ Taktéž bylo během analýzy zjištěno, že mnohé z bankrotujících firem nevykazovaly vůbec žádné tržby, čímž byl znemožněn výpočet některých ukazatelů (doba obratu zásob apod.). Proto bylo rozhodnuto tyto firmy (celkem 28) ze vzorku vyloučit.

3.1 Jednofaktorová analýza a úprava vstupních dat

Nejprve byla provedena jednofaktorová analýza, jejíž pomocí byly určeny ty ukazatele, jež jsou pro predikci bankrotu statisticky významné. V následující tabulce 1 jsou uvedeny hodnoty pro analýzu vlivů jednotlivých vysvětlujících veličin na bankrot firmy. V tabulce jsou zobrazeny p-hodnoty tzv. z-statistiky, jejichž hodnota musí být menší než 0,05, aby vliv zkoumané proměnné byl statisticky významný.

Po analyzování statistické významnosti vlivu jednotlivých veličin bylo vyloučeno celkem 11 ukazatelů. Jako nevhodné pro predikci byly určeny všechny tři ukazatele reprezentující oblast aktivity, čili doba obratu zásob, doba obratu pohledávek a doba obratu závazků. Taktéž ukazatel finanční páky, úvěrové zadluženosti, úrokového zatížení a ukazatel zadluženosti vlastního kapitálu byly statisticky nevýznamné. Vyřazen byl rovněž ukazatel ROA, který vycházel z nezdáněného zisku před odpisy, úroky a daněmi (EBITDA), přestože byl statisticky významný. Důvodem vyřazení byl obdobný princip výpočtu ukazatele ROA, který vychází ze zisku před úroky a daněmi

¹ Např. kladná rentabilita vlastního kapitálu při záporném výsledku hospodaření a záporném vlastním kapitálu z důvodu akumulace dlouhodobých ztrát.

Tabulka 1 Statistická významnost vlivu daného ukazatele na bankrot firmy

	Ukazatel		Výpočet	P> z
Vysvětlující veličiny	Rentabilita	ROA (EBITDA)	$\frac{EBITDA}{aktiva}$	0,000
		ROA	$\frac{EBIT}{aktiva}$	0,000
		ROS	$\frac{EAT}{tržby}$	0,003
		Přidaná hodnota/tržby	$\frac{PH}{Tržby}$	0,000
	Zadluženost	Celková zadluženost	$\frac{cizí\ kapitál}{aktiva}$	0,000
		Běžná zadluženost	$\frac{krátkodobý\ cizí\ kapitál}{aktiva}$	0,000
		Zadluženost vlastního kapitálu	$\frac{cizí\ kapitál}{vlastní\ kapitál}$	0,952
		Úvěrová zadluženost	$\frac{úvěry}{vlastní\ kapitál}$	0,501
		Úvěry/pasiva	$\frac{bankovní\ výpomoci}{pasiva}$	0,000
		Finanční páka	$\frac{aktiva}{vlastní\ kapitál}$	0,946
		Úrokové zatížení	$\frac{úroky}{EBIT}$	0,166
		Stupeň krytí stálých aktiv	$\frac{dlouhodobý\ kapitál}{stálá\ aktiva}$	0,017
	Aktivita	Doba obratu zásob	$\frac{zásoby \cdot 360}{tržby}$	0,093
		Doba obratu pohledávek	$\frac{pohledávky \cdot 360}{tržby}$	0,221
		Doba obratu závazků	$\frac{závazky \cdot 360}{tržby}$	0,267
	Likvidita	Běžná likvidita	$\frac{oběžný\ majetek}{krátkodobé\ závazky}$	0,000
		Pohotová likvidita	$\frac{OA - zásoby}{krátkodobé\ závazky}$	0,000
		Okamžitá likvidita	$\frac{pohotové\ peníze\ a\ ekvivalenty}{okamžitě\ splatné\ dluhy}$	0,000
		Poměrový ukazatel likvidity	$\frac{OA - krátkodobé\ závazky}{OA}$	0,004

(EBIT). Pro vytvoření modelu by dva velmi podobné ukazatele byly zbytečné. Také byl vyloučen ukazatel běžné likvidity a poměrový ukazatel likvidity, neboť oblast likvidity dostatečně reprezentují zbylé dva ukazatele, tj. okamžitá a pohotová likvidita. Ukazatel stupně krytí stálých aktiv byl rovněž vypuštěn, a to z důvodu, že v daném vzorku firem bylo hned několik desítek společností s nulovým dlouhodobým majetkem.

Dalším krokem u zbylých osmi ukazatelů bylo odstranění extrémních² hodnot. K vypuštění firem s nesmyslnými či velmi extrémními hodnotami bylo přistoupeno z toho důvodu, aby ve vzorku nebyly zahrnuty takové firmy, které již evidentně trpí příznaky finančních potíží a dosud na ně nebyl prohlášen konkurz, zahájeno insolvenční řízení, apod. Extrémní hodnoty, které byly i tak smysluplné, je však možné ve vzorku ponechat a zvolit vhodnou metodu transformace těchto vysvětlujících veličin. Lze uvést například použití kumulativní distribuční funkce normálního rozdělení, Royston a Sauerbrei (2007), nebo negativní exponenciální transformaci, Sauerbrei a Royston (1999). Výsledky jednofaktorových logitových modelů po odstranění firem s extrémními hodnotami jsou zachyceny v tabulce 1 a znázornění výsledného vztahu pravděpodobnosti bankrotu a zvolených nezávislých proměnných je znázorněn na obrázku 1.

Na základě tohoto obrázku lze konstatovat, že s rostoucí rentabilitou aktiv dochází ke snížení pravděpodobnosti vzniku bankrotu. Pokud tvoří provozní zisk přibližně 50 % aktiv, je pravděpodobnost finanční tísně téměř nulová. Obdobnou závislost nalezneme i v případě rentability tržeb. Se zvyšující se rentabilitou klesá pravděpodobnost problémů v hospodaření. Obdobně je tomu u ukazatele pohotové likvidity. Obecně doporučovaná hranice od 1,0 do 1,5 zde představuje relativně vysoké riziko bankrotu (do 50 %). Obdobně je tomu u okamžité likvidity, kdy u doporučené výše 0,2 je riziko úpadku cca 55 %, jak

² Za extrémní hodnoty byly označeny ty, jež se nacházely vně intervalu, jehož hranice byly určeny tak, že od dolního kvantilu byl odečten trojnásobek rozdílu třetího a prvního kvantilu, respektive ke třetímu kvantilu byl přičten trojnásobek rozdílu třetího a prvního kvantilu.

Tabulka 2 Model odhadnutý pomocí postupné regrese

Log likelihood = -108.30708				Pseudo R2	=	0.3054
bankrot	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
roa	-3.220297	1.282398	-2.51	0.012	-5.733751	- .7068438
ok_lik	-1.876355	.5129237	-3.66	0.000	-2.881667	- .8710426
celk_zad1	1.915746	.5395233	3.55	0.000	.8583001	2.973193
uv_pas	1.977485	1.162686	1.70	0.089	-.3013369	4.256307
phtrczy	1.010761	.877616	1.15	0.249	-.7093349	2.730856
_cons	-1.588707	.5406132	-2.94	0.003	-2.648289	-.5291246

Zdroj: zpracování v programu STATA 10.1

dokládá graf. Avšak i zde je závislost nepřímá, tj. s rostoucím ukazatelem nebezpečí bankrotu klesá. Rovněž z grafu vyplývá, že s rostoucím podílem přidané hodnoty na tržbách klesá riziko bankrotu. V případě, že přidaná hodnota tvoří přibližně 30 % tržeb, je riziko krachu nižší než 40 %.

Naopak je tomu u ukazatele celkové zadluženosti a poměru krátkodobého kapitálu k aktivům, kdy s rostoucí úrovní těchto ukazatelů roste i riziko bankrotu. Optimální hranice podílu úvěrů na pasivech se dle grafu zdá být do 20 %.

3.2 Vícefaktorová analýza

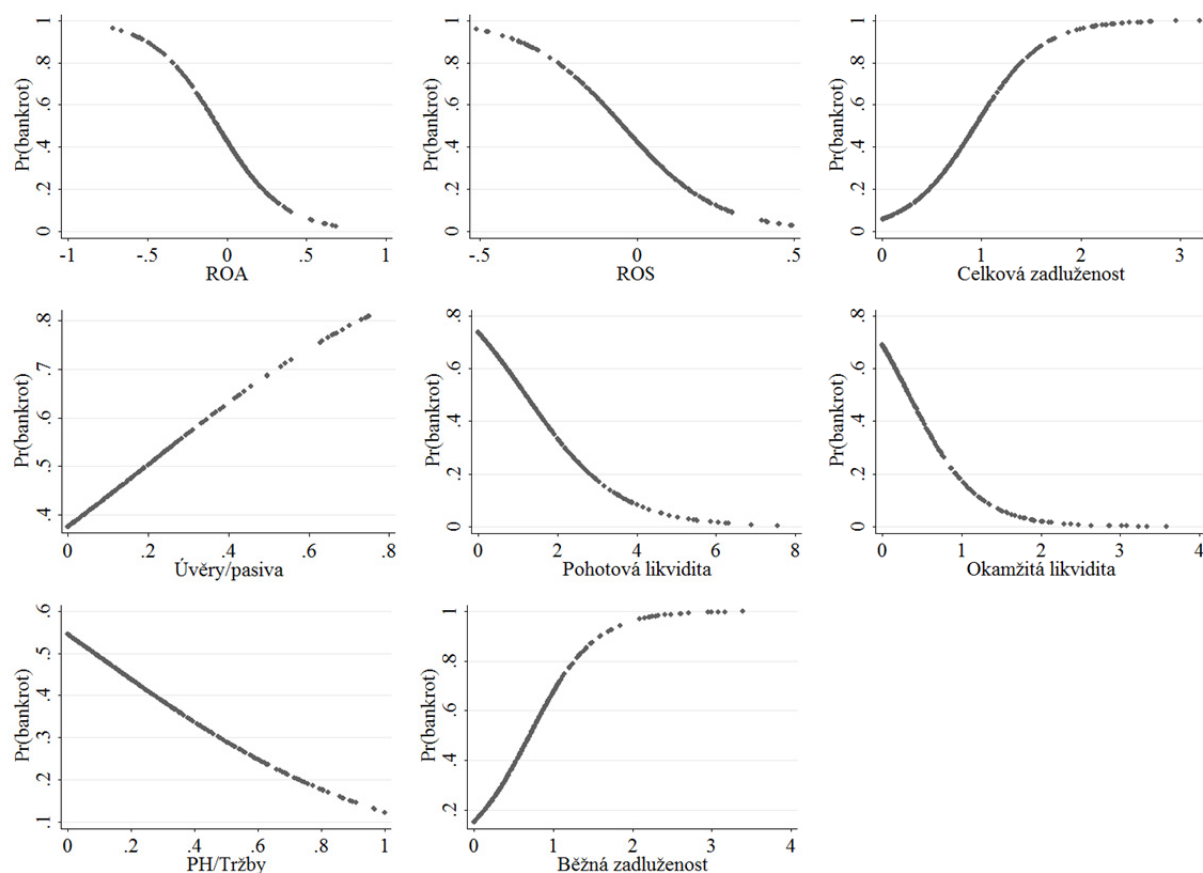
Po vícefaktorové analýze, na jejímž základě bylo vyloučeno 11 poměrových ukazatelů, a po zjištění vlivu nezávisle proměnné na hlavní vysvětlovanou veličinu následovalo vytvoření modelu se všemi proměnnými metodou postupné regrese. Hranice statistické významnosti (*p*-hodnota) byla zvolena na úrovni 25 %, aby nedošlo k opomenutí některé z významnějších proměnných. Výsledky vícefaktorové analýzy jsou zachyceny v tabulce 2.

Z této tabulky je patrné, že po zahrnutí všech veličin do modelu se statisticky nevýznamným stal ukazatel rentability tržeb, ukazatel pohotové likvidity a ukazatel celkové zadluženosti. Mezi těmito třemi ukazateli a ostatními existují zřejmě interakce, jejichž působením se stávají jiné faktory statisticky nevýznamnými. I přes toto zjištění byly interakce zanedbány a k predikci úpadku firmy bylo použito zbylých pět poměrových ukazatelů. Avšak i po vyloučení těchto dalších ukazatelů zůstává zastoupena jak oblast rentability, tak likvidity a zadluženosti.

3.3 Ověření linearity logitu

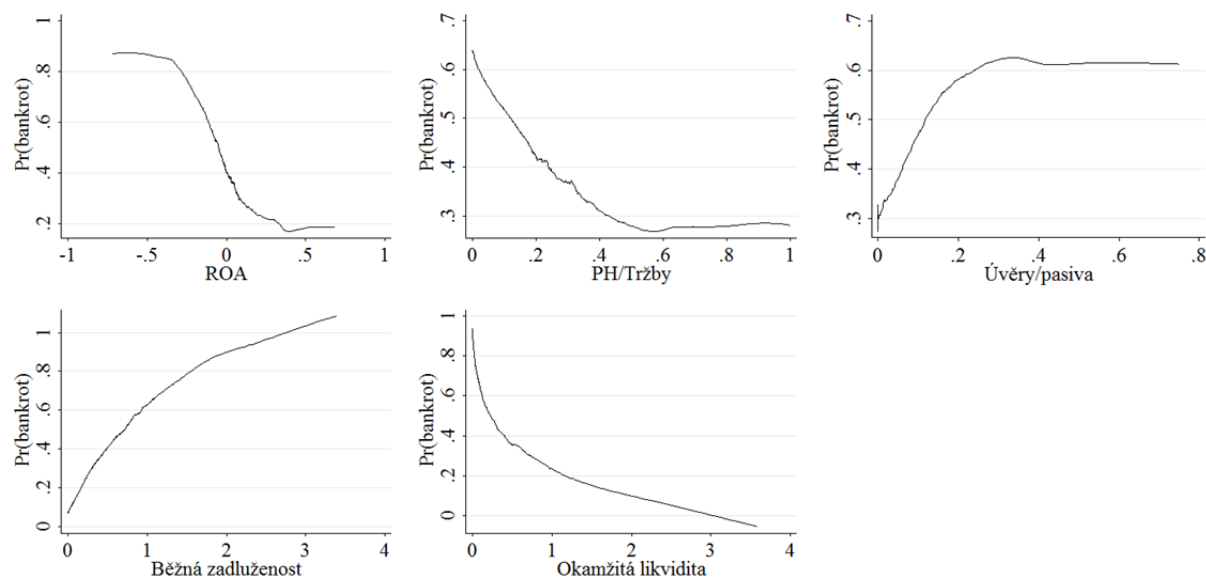
Dalším krokem bylo ověření, zda vliv jednotlivých veličin lze vyjádřit lineární funkcí, tedy zda logit (5) je lineární. Toto bylo provedeno nejprve graficky pomocí vyhlazeného logitu metodou *lowess*, viz obrázek 2. Poté byly všechny ukazatele nahrazeny částečným polynomem druhého řádu a statisticky otestovány.

Na obrázku 2 jsou znázorněny grafy, které vyjadřují závislost logitu na dané vysvětlující veličině. Jak již bylo zmíněno, v případě, že funkce logitu není lineární, je třeba vztah *linearizovat*. Z obrázku je sice



Obrázek 1 Pravděpodobnost vzniku bankrotu dle hodnoty daného ukazatele

Zdroj: zpracování v programu STATA 10.1



Obrázek 2 Grafické testování linearity logitů

Zdroj: zpracování v programu STATA 10.1

zřejmé, že některé vztahy lze dobře aproximovat přímkou (tedy lineárním vztahem), avšak u jiných není tento závěr tak jednoznačný. Proto bylo o linearity, respektive nelinearitě logitů rozhodnuto na základě testování částečného polynomu druhého řádu (FP2) věrohodnostním poměrem. Výsledky jsou uvedeny v následující tabulce 3.

Tabulka 3 Testy věrohodnostním poměrem lineárních a nelineárních logitů

Ukazatel	P (*)		
	Nezařazen	Lineární logit	FP1
ROA	0,000	0,055	0,022
PH/tržby	0,000	0,055	0,253
Úvěry/pasiva	0,000	0,006	0,690
Běžná zadluženost	0,000	0,127	0,338
Okamžitá likvidita	0,000	0,000	0,003

Testováno bylo zařazení ukazatele do modelu, významnost lineárního logitu a linearizovaného logitu pomocí částečného polynomu prvního řádu (FP1) proti částečného polynomu řádu 2 (FP2). *P*-hodnoty v tabulce 3 jsou výsledky testu věrohodnostních poměrů mezi daným modelem a modelem FP2. Pokud je *p*-hodnota nižší než 0,05, pak model FP2 dává lepší výsledek neboli nižší věrohodnostní poměr než testovaný model a vybrán je model FP2.

Dle výsledků lze tedy konstatovat, že u dvou z pěti vysvětlujících proměnných bylo pomocí částečného polynomu zjištěno, že jejich vliv na bankrot nelze

výjádřit jako lineární funkci a bylo nutné je převést na veličiny kategoriální (jedná se o celkovou zadluženost a okamžitou likviditu). K tomuto účelu byly obě veličiny kategorizovány na čtyři kategorie pomocí kvartilů.

Dále byl tedy vytvořen nový model, který zahrnoval kategorizované veličiny a ostatní původní spojité veličiny, tedy rentabilitu aktiv, podíl přidané hodnoty a tržeb a běžnou zadluženost. Odhadnutý model je zachycen v následující tabulce 4.

Z tabulky 4 je patrné, že při zahrnutí všech veličin se jako statisticky nevýznamný na základě *z*-statistiky jeví ukazatel PH/tržby. Z tohoto důvodu byl v dalším kroku vytvořen nový model predikce úpadku (2), a to bez zahrnutí této statisticky nevýznamné veličiny. Výsledek je zachycen v tabulce 5.

Z tabulky 5 je zřejmé, že k výraznějším změnám ve statistické významnosti vysvětlujících proměnných nedošlo. Zhoršila se nepatrně významnost pouze u druhé kategorie ukazatele úvěry/pasiva, přesto lze tuto kategorii považovat za významnou, a proto byla v modelu ponechána. Na základě těchto výsledků však nebylo možné rozhodnout, zda je druhý model lepší než model předchozí, proto byl druhý model otestován vůči prvnímu věrohodnostním poměrem. Výsledek je uveden v tabulce 6.

Test věrohodnostním poměrem porovnává první (složitější) model vůči druhému (jednoduššímu) modelu. Sledována je přitom hodnota označená jako Prob > chi2, která vyjadřuje statistickou významnost

Tabulka 4 Model predikce úpadku (1)

Log likelihood = -108.47201			Pseudo R2 = 0.3375		
bankrot	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
roa	-2.891227	1.199347	-2.41	0.016	-5.241903 - .5405511
bez_zadl	1.904589	.5018027	3.80	0.000	.9210735 2.888104
phtrzby	1.381506	.8599234	1.61	0.108	-.3039131 3.066925
uv_pas2	.9535791	.4347452	2.19	0.028	.1014941 1.805664
uv_pas3	1.195497	.4387816	2.72	0.006	.335501 2.055493
ok_lik2	-1.476553	.3962488	-3.73	0.000	-2.253186 -.6999192
ok_lik3	-2.050313	.5460406	-3.75	0.000	-3.120533 -.9800928
_cons	-1.493461	.5662144	-2.64	0.008	-2.603221 -.3837014

Zdroj: zpracování v programu STATA 10.1

Tabulka 5 Model predikce úpadku (2)

Log likelihood = -109.77858			Pseudo R2 = 0.3295		
bankrot	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
roa	-2.553049	1.160016	-2.20	0.028	-4.82664 - .2794594
bez_zadl	1.770407	.4751643	3.73	0.000	.8391019 2.701712
uv_pas2	.9337767	.4346423	2.15	0.032	.0818934 1.78566
uv_pas3	1.070163	.428336	2.50	0.012	.2306397 1.909686
ok_lik2	-1.427431	.3894382	-3.67	0.000	-2.190716 -.6641464
ok_lik3	-1.87587	.5287662	-3.55	0.000	-2.912233 -.8395076
_cons	-1.060944	.4840182	-2.19	0.028	-2.009603 -.112286

Zdroj: zpracování v programu STATA 10.1

Tabulka 6 Srovnání modelu predikce úpadku (1) a (2)

Likelihood-ratio test (Assumption: χ^2 nested in model 1)		LR chi2(1) = 2.61
		Prob > chi2 = 0.1060

Zdroj: zpracování v programu STATA 10.1

prvního modelu vůči druhému. Vzhledem k tomu, že tato hodnota je vyšší než 0,05, lze v tomto případě říci, že první model není statisticky významnější než druhý, a pro predikci úpadku jsou tak oba modely považovány za ekvivalentní, tedy že rozdíl mezi těmito dvěma modely není statisticky významný. Na základě těchto skutečností se dále pracovalo se zjednodušeným modelem predikce úpadku s vynecháním ukazatele PH/tržby, tedy s modelem predikce úpadku (2).

3.4 Vyhodnocení modelu

Ke zhodnocení spolehlivosti při predikování finanční tísně firmy modelem predikce úpadku (2) posloužily výsledky uvedené v tabulce 7.

Na základě hodnot klasifikační tabulky můžeme o testovaném modelu bankrotu prohlásit, že jeho úspěšnost diskriminace je 79,34 %. V 75 z 99 případů (v 75,76 %), kdy došlo k úpadku společnosti, model tuto situaci identifikoval správně, ve zbylých 24 případech (24,24 %) tuto situaci odhadl chybně. V 143 případech, kdy společnost nezkrachovala, odhadl tuto skutečnost v 117 případech správně (81,82 %), mylil se v 26 případech (18,18 %).

Tabulka 7 Klasifikační tabulka

Classified	True		Total
	D	~D	
+	75	26	101
-	24	117	141
Total	99	143	242

Classified + if predicted $\Pr(D) \geq .5$
True D defined as bankrot != 0

Sensitivity	$\Pr(+ D)$	75.76%
Specificity	$\Pr(- \sim D)$	81.82%
Positive predictive value	$\Pr(D +)$	74.26%
Negative predictive value	$\Pr(\sim D -)$	82.98%
False + rate for true ~D	$\Pr(+ \sim D)$	18.18%
False - rate for true D	$\Pr(- D)$	24.24%
False + rate for classified +	$\Pr(\sim D +)$	25.74%
False - rate for classified -	$\Pr(D -)$	17.02%
Correctly classified		79.34%

Zdroj: zpracování v programu STATA 10.1

Grafickým znázorněním klasifikační tabulky je následující obrázek 3, který zachycuje vztah mezi senzitivitou a specifitou.

V pravé části obrázku 3 je zachycena zkonstruovaná křivka ROC, která vyjadřuje vztah mezi senzitivitou a hodnotou 1-specifita.

Křivka ROC se na obrázku 3 přibližuje levému hornímu rohu, což značí vysokou spolehlivost modelu při klasifikaci společností. Pro hodnocení modelu je vhodné použít i ukazatel AUC, jehož hodnota činí 0,8625, což se velmi blíží ideální hodnotě, kterou je hodnota 1.

3.5 Srovnání modelu

Na závěr srovnáme vybrané studie, které byly provedeny v oblasti scoringových modelů úpadku firem, a to zejména se studií na bázi logitu, Jakubík a Teplý (2011), z let 1993–2005, a dále s vybranými modely na bázi vícefaktorové diskriminační analýzy za období 1971–1995, Altman (1996), a z roku 2004, Neumaierová a Neumaier (2005). V následující tabulce 8 jsou tyto modely shrnuty včetně označení, zda vybrané ukazatele byly zahrnuty do modelu. Před samotným srovnáním je vhodné poznamenat, že do prvotní jednofaktorové analýzy nevstupovaly stejné ukazatele jako do ostatních studií. Ve srovnání s Jakubíkem a Teplým (2011) zde nebyl zahrnut ukazatel ROE a dále došlo k drobným odlišnostem v podobě ukazatele ROS (zde počítán s čistým ziskem místo provozního) a úrokového krytí (zde použita obrácená hodnota, tedy úrokové zatížení).

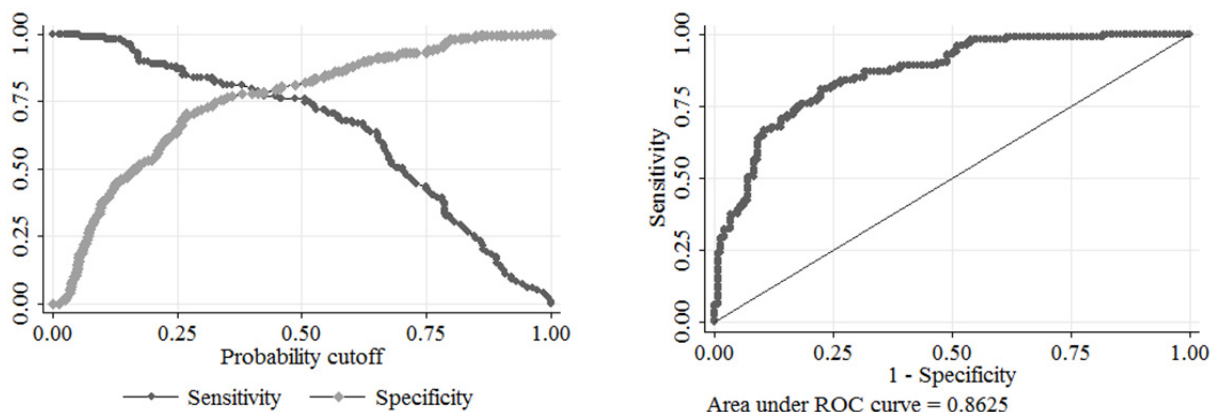
Pro predikci úpadku českých firem v roce 2009 se ukázala statisticky významná výše ukazatele: (1) *podíl úvěrů na pasivech*, (2) *celková zadluženost*, (3) *okamžitá likvidita* a (4) *rentabilita aktiv* (ROA). Ostatní ukazatele nebyly do modelu zahrnuty. Zůstává otázkou, zda tento model je vhodnější pro krátký horizont predikce, daný rok 2009 nebo je výsledek způsoben použitým (úzkým) vzorkem. Nicméně jiné ukazatele vybrané v dlouhodobém modelu Jakubíka a Teplého (2011) nejsou zde vůbec statisticky významné, a to (1) *podíl úvěrů a vlastního kapitálu*, (2) *úrokové krytí*, (3) *rentabilita tržeb* a (4) *dooba obratu zásob*.

Tabulka 8 Srovnání modelu s jinými studiemi

	Tato studie (2012) Logit	Jakubík, Teplý (2011) Logit	Neumaierová a Neumaier (2005) Více.disk. analýza	Altman (1996) Více.disk. analýza
Úvěry/pasiva	ano			
Celková zadluženost	ano			
Okamžitá likvidita	ano	ano		ano
ROA	ano		ano	

Zdroj: Jakubík a Teplý (2011)

Ukazatelem zahrnutým do dvou ostatních zmíněných studií je ukazatel *okamžitá likvidita* a ukazatel *ROA* do modelu dle Neumaierová a Neumaier (2005). Přestože ostatní ukazatele byly zahrnuty pouze v této studii, lze obecně říci, že i zde jsou zastoupeny oblasti zadluženosti i rentability. Lze tedy konstatovat, že i zde bylo dosaženo velmi obdobného modelu. Za největší odlišnost však lze považovat absenci ukazatele reprezentující oblast aktivity. Zda lze tyto ukazatele



Obrázek 3 Vztah mezi senzitivitou a specifivitou výsledného modelu (vlevo) a křivka ROC (vpravo)

Zdroj: zpracování v programu STATA 10.1

chápat také jako krátkodobé determinanty úpadku, by mělo být však podrobeno dalšímu zkoumání.

4. Závěr

Příspěvek byl věnován scoringovému modelu úpadku na bázi logistické regrese. Tento model byl vytvořen tak, aby co nejlépe klasifikoval *dobré* a *špatné* firmy. K tomuto účelu byla provedena finanční analýza 400 českých firem a z vypočtených ukazatelů byly vybrány ty, které měly největší klasifikační schopnost. Nejprve byla provedena jednofaktorová analýza vlivu jednotlivých ukazatelů na pravděpodobnost úpadku a poté byl zkonstruován model obsahující množinu ukazatelů, které mají největší klasifikační schopnost. Celý model byl následně statisticky verifikován a validace modelu byla provedena klasifikační tabulkou a ROC křivkou. Vzhledem k tomu, že model byl odhadnut na krátké časové řadě (v rámci jednoho roku), mohou být koeficienty modelu nestabilní v čase, z čehož vyplývá omezení modelu. Jeho použití je tedy vhodné pouze pro krátkodobý časový horizont (po roce 2008) nebo ve stejné fázi ekonomického cyklu.

Z jednofaktorové analýzy mj. vyplynulo, že ukazatele ROA, ROS, pohotová a okamžitá likvidita a podíl přidané hodnoty na tržbách mají pozitivní vliv na úpadek (s rostoucí hodnotou pravděpodobnost úpadku klesá) a naopak celková a běžná zadluženost mají negativní vliv. Při ověření linearity pomocí metody lowess a částečných polynomů bylo zjištěno, že vliv celkové zadluženosti a okamžité likvidity je nelineární. Tato nelinearita byla vyřešena tak, že hodnoty ukazatelů byly *kategorizovány* do několika skupin. Do konečného modelu byly nakonec na základě statistické významnosti vybrány pouze ukazatel ROA, celková zadluženost, podíl úvěrů a pasív a okamžitá likvidita. I přes vyloučení celé řady jiných ukazatelů zůstala zastoupena oblast rentability, zadluženosti a likvidity.

Z následné validace modelu byla následně zjištěna velmi vysoká schopnost klasifikace, tedy 79 %.

Navržený model má však také některé nedostatky, které lze odstranit a celý model tak ještě rozšířit a zpřesnit podle toho, k jakému účelu by měl sloužit. Především by měla být provedena dodatečné ověření, zda opravdu nelze nalézt vhodný ukazatel aktivity, který by byl pro model úpadku statisticky významný. Dále nelineární vliv ukazatelů je vhodné ponechat jako spojitou veličinu a nekategorizovat ji, čímž se získá lepší vztah pro kvantifikaci vlivu těchto determinant. Na druhou stranu lze vyzdvihnout, že v modelu byl, naproti jiným studiím, částečně ošetřen nelineární vliv proměnných pomocí jejich kategorizace. Další možné rozšíření vyplynulo ze zjištěných interakcí, které mohou být zanedbány pro účely klasifikace či predikce, ale pro analýzu vlivu daného ukazatele je nutné tyto interakce rovněž modelovat. Neposledním rozšířením je také externí validace modelu a rozšíření datové základny o další firmy nebo doplnění finanční analýzy o kvalitativní či subjektivní hodnocení.

Literatura

- AI, C., NORTON, E. (2003). Interaction terms in logit and probit models. *Economics Letters* 80: 123–129. [http://dx.doi.org/10.1016/S0165-1765\(03\)00032-6](http://dx.doi.org/10.1016/S0165-1765(03)00032-6)
- ALTMAN, E. (1996). Corporate bonds and commercial loan portfolio analysis. *Risk Management in Banking*, The Wharton Financial Institutions Center.
- ALTMAN, E. (1968). Financial ratios, discriminate analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23(4): 589–609. <http://dx.doi.org/10.2307/2978933>

- ALTMAN, E. a kol. (2010). *The Z-Metrics Methodology For Estimating Company Credit Ratings and Default Risk Probabilities*. New York: RiskMetrics Group.
- ALTMAN, E., HOTCHKISS, E. (2006). *Corporate Financial Distress and Bankruptcy*. Hoboken: Wiley.
- BEAVER, W. (1967). Financial ratios as predictors of failures. *Empirical Research in Accounting: Selected Studies, Journal of Accounting Research* 5: 71–102.
- GREENE, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters* 107: 291–296.
<http://dx.doi.org/10.1016/j.econlet.2010.02.014>
- GURNÝ, P. (2011). The stability investigation of the three large Czech banks within Z-metrics methodology. In: *International Conference on Mathematical Methods in Economics 2011*. Prague: University of Economics, 205–211.
- GURNÝ, P., GURNÝ, M. (2010a). Logit and probit model within estimation of US banks. In: *Proceedings of 47th EWGFM Meeting*. Ostrava: VŠB-TUO, 73–80.
- GURNÝ, P., GURNÝ, M. (2010b). Comparison of the credit scoring models on PD estimation of US banks. In: *28th International Conference on Mathematical Methods in Economics 2010*. Ceske Budejovice: University of West Bohemia, 188–194.
- HEBÁK, P. a kol. (2007). *Vícerozměrné statistické metody III*. Praha: Informatorium.
- HEBÁK, P. a kol. (2004). *Vícerozměrné statistické metody I*. Praha: Informatorium.
- HOSMER, D. W., LEMESHOW, S. (2000). *Applied Logistic Regression*. New York: Wiley-Interscience.
<http://dx.doi.org/10.1002/0471722146>
- JAKUBÍK, P. (2007). Macroeconomic environment and credit risk. *Czech Journal of Economics and Finance* 57(1–2): 60–78.
- JAKUBÍK, P., TEPLÝ, P. (2011). The JT index as an indicator of financial stability of corporate sector. *Prague Economic Papers* 2: 157–176.
- JAKUBÍK, P., TEPLÝ, P. (2008). The prediction of corporate bankruptcy and Czech economy financial stability through logit analysis. *Institute of Economic Studies Working Paper*, No. 19. Prague: IES FSV, Charles University.
- MCFADDEN, D. (1976). A comment on discriminant analysis versus logit analysis. *Annals of Economic and Social Measurement* 5 (4): 511–523.
- NEUMAIEROVÁ, I., NEUMAIER, I. (2005). Index IN05. In: *Evropské finanční systémy: Sborník příspěvků z mezinárodní vědecké konference*. Brno: Masarykova univerzita v Brně, 143–146.
- OHLSON, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *The Journal of Accounting Research* 18(1): 109–131.
<http://dx.doi.org/10.2307/2490395>
- PECÁKOVÁ, I. (2007). Logistická regrese s vícekategoriální vysvětlovanou proměnnou. *Acta Oeconomica Pragensia* 1: 86–86.
- ROYSTON, P., ALTMAN, D. G. (1994). Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling (with discussion). *Applied Statistics* 43: 429–467.
<http://dx.doi.org/10.2307/2986270>
- ROYSTON, P., SAUERBREI, W. (2007). Improving the robustness of fractional polynomial models by preliminary covariate transformation. *Computational Statistics and Data Analysis* 51: 4240–4253.
<http://dx.doi.org/10.1016/j.csda.2006.05.006>
- SAUERBREI, W., ROYSTON, P. (1999). Building multivariable prognostic and diagnostic models: transformation of the predictors using fractional polynomials. *Journal of the Royal Statistical Society, Series A* 162: 71–94. <http://dx.doi.org/10.1111/1467-985X.00122>
- SEDLÁČEK, J. (2007). *Finanční analýza podniku: Metody finanční analýzy, Horizontální, vertikální a indexová analýza, Bonitní a bankrotní modely, Cash flow a rating*. Brno: Computer Press.